**MAT 2379, Introduction to biostatistics**

**Solution to Assignment 2**
(Total = 32 points)

**Problem 3.18** Let $A$ be the event that the child is diagnosed with autism before the age of six and $B$ the event that the child has been vaccinated against measles prior to the age of two. We know that

$$P(A|B) = \frac{38}{2565} = 0.015.$$

Since this value coincides with $P(A)$, we conclude that vaccination against measles and autism are independent.

**Problem 4.10**

**[2]** (a) Since the total probability is equal to 1, then

$$1 = \frac{c}{18} + \frac{c}{18} + \frac{2c}{9} + \frac{c}{6} + \frac{c}{3} + \frac{c}{3} + \frac{c}{6} = \frac{4\,c}{3}.$$

Thus, $c = 3/4$.

**[4]** (b) The probability mass function is

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $f(x)$ | 1/24 | 1/24 | 1/6 | 1/8 | 1/4 | 1/4 | 1/8 |

The expected number of properly stained cells is

$$E(X) = 0\,(1/24) + 1\,(1/24) + \cdots + 6\,(1/8) = 3.75.$$

**[6]** (c) $P(|X-2| \leq 1) = P(X=1) + P(X=2) + P(X=3) = 1/3 = 0.3333$ Using $\mu = E(X) = 3.75$, we compute

$$\sigma = \sqrt{0^2\,(1/24) + 1^2\,(1/24) + \cdots + 6^2\,(1/8) - (3.75)^2} = 1.5877.$$

Thus,

$$P(X > \mu + \sigma) = P(X > 5.3377) = P(X = 6) = \frac{1}{8} = 0.125.$$

**Problem 4.14** Let $X$ be the number of defective packages among $n = 25$ packages. $X$ has a binomial distribution with $n = 20$ and $p = 0.05$.
(a) We want $P(X \geq 1) = 1 - P(X = 0) = 1 - (0.95)^{25} = 0.7226$.
(b) The expected number of defective packages is $E[X] = n\,p = 1.25$.
(c) Let $X$ be the number of defective packages among $n$ packages. $X$ has a binomial distribution with $n$ trials and $p = 0.05$. We want $P(X \geq 1) \geq 0.9$, but $P(X \geq 1) = 1 - P(X = 0) = 1 - (0.95)^n$. Solving $1 - (0.95)^n \geq 0.9$ for $n$ gives

$$n \geq \frac{\ln(0.1)}{\ln(0.95)} = 44.89.$$

We should select at least 45 packages.

**Problem 5.4** Let $X$ be the size of the body (in cm) of a snowy owl. $X$ has a normal distribution with $\mu = 61.5$ and $\sigma = 4.75$.

**[2]** (a) We want

$$P(X > 65) = 1 - P\left(Z \leq \frac{65 - 61.5}{4.75}\right) = 1 - \Phi(0.74) = 0.2296.$$

**[2]** (b) We want

$$P(X < 55) = P\left(Z < \frac{55 - 61.5}{4.75}\right) = \Phi(-1.37) = 0.0853.$$

**[4]** (c) Let $Y$ be the number of snowy owls in the sample with a body greater than 55 cm. $Y$ has a binomial distribution with $n = 10$ and $p = P(X > 55) = 1 - 0.0853 = 0.9147$. We want $E[Y] = n\,p = 10\,(0.9147) = 9.147$.

Let $W$ be the number of snowy owls in the sample with a body greater than 65 cm. $W$ has a binomial distribution with $n = 10$ and $p = P(X > 65) = 0.2296$. We want

$$P(W \leq 1) = (1-p)^{10} + 10\,p\,(1-p)^9 = 0.2931.$$

**[2]** (d) Let $X$ be the size of the body (in cm) of a snowy owl. $X$ has a normal distribution with $\mu = 61.5$ and $\sigma$ is unknown. Since, $0.25 = P(X > 65)$ or equivalently

$$0.75 = P(X < 65) = P\left(Z < \frac{65 - 61.5}{\sigma}\right),$$

then, $0.675 = (65 - 61.5)/\sigma$. (We used the fact that $\Phi(0.67) = 0.7486$ and $\Phi(0.68) = 0.7517$.) Therefore, $\sigma = (65 - 61.5)/0.675 = 5.1851$.

**[2]** **Additional question**: Let $I_i$ be the event that the $i$-th item identifies correctly the disease, $i = 1, 2, 3$. So $P(I_1) = 0.88$, $P(I_2) = 0.90$, and $P(I_3) = 0.95$. Since items are independent, we have

$$P(I_1 \cup I_2 \cup I_3) = 1 - P(I_1' \cap I_2' \cap I_3') = 1 - (0.12) \times (0.10) \times (0.05) = 0.9994.$$

Or, equivalently, we have

$$\begin{aligned} P(I_1 \cup I_2 \cup I_3) &= P(I_1) + P(I_2) + P(I_3) - P(I_1 \cap I_2) - P(I_1 \cap I_3) - P(I_2 \cap I_3) + P(I_1 \cap I_2 \cap I_3) \\ &= 0.88 + 0.90 + 0.95 - 0.88 \times 0.90 - 0.88 \times 0.95 - 0.90 \times 0.95 \\ &+ 0.88 \times 0.90 \times 0.95 = 0.9994. \end{aligned}$$

**Part (II)**

1. $X \sim Bin(25, 0.4)$, thus we have

**[2]** **a)** $P(X \leq 10) = 0.5858$;

```
> pbinom(10, size=25, prob=0.4)
[1] 0.585775
```

**[2]** **b)** $P(X \geq 15) = 1 - P(X \leq 14) = 0.0344$;

```
> 1-pbinom(14, size=25, prob=0.4)
[1] 0.03439152
```

**[2]** **c)** $P(X < 7) = P(X \leq 6) = 0.0736$;

```
> pbinom(6, size=25, prob=0.4)
[1] 0.07356526
```

**[2]** **d)** $P(12 \leq X < 20) = P(X \leq 19) - P(X \leq 11) = 0.2677$.

```
> pbinom(19, size=25, prob=0.4)-pbinom(11, size=25, prob=0.4)
[1] 0.2676642
```

2. $X \sim N(12, 9)$, thus we have

**a)** $P(X < 9) = 0.5328$;

```
> pnorm(9, mean=12, sd=3)
[1] 0.1586553
```

**b)** $P(10.5 < X < 15) = P(X < 15) - P(X < 10.5) = 0.5328$.

```
> pnorm(15, mean=12, sd=3)-pnorm(10.5, mean=12, sd=3)
[1] 0.5328072
```

**c)** We want $x_0$ such that $P(X < x_0) = 0.95$.
Using `qnorm` function, we find $x_0 = 16.93456$.

```
> qnorm(0.95, mean=12, sd=3)
[1] 16.93456
```

**d)** We want $x_0$ such that $P(|X - 12| > 3x_0) = 0.2$.

The statement is equivalent to find a $x_0$ such that $0.2 = P(\frac{|X-12|}{3} > x_0) = P(|Z| > x_0) = P(Z > x_0) + P(Z < -x_0)$. Note that the PDF of $Z$ is symmetric about 0; this implies that

$$P(Z < -x_0) = P(Z > x_0) = (0.2)/2 = 0.1.$$

So we want $x_0$ such that $P(Z < -x_0) = 0.1$. Using R, we find $-x_0 = -1.281552$, which implies $x_0 = 1.281552$.

```
> qnorm(0.1, mean=0, sd=1)
[1] -1.281552
```